



ZW  
AR

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE  
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

In re Application of

Atty. Docket

ERIC COHEN-SOLAL ET AL

US000015

Serial No. 09/488,028

Group Art Unit: 2173

Filed: JANUARY 20, 2000

Examiner: T. HAILU

Conf. # 3849

TITLE: **MULTI-MODAL VIDEO TARGET ACQUISITION AND RE-DIRECTION  
SYSTEM AND METHOD**

Mail Stop Appeal Brief - Patents  
Commissioner for Patents  
PO Box 1450  
Alexandria, VA 22313-1450

Enclosed is an Appeal Brief in the above-identified  
patent application.

The fee for this Appeal Brief submission was previously  
paid in connection with an Appeal Brief submitted on June 7, 2004.  
No further fee is believed due at this time.

---

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited this date  
with the United States Postal Service as first-class mail in an envelope addressed to:

Mail Stop Appeal Brief - Patents  
COMMISSIONER FOR PATENTS  
PO Box 1450  
Alexandria, VA 22313-1450

On September 15, 2005

By Gregory L. Thorne  
Gregory L. Thorne



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Before the Board of Patent Appeals and Interferences

In re the Application

Inventor : ERIC COHEN-SOLAL ET AL  
Application No.: 09/488,028  
Filed : JANUARY 20, 2000  
Conf. No. : 3849  
For : MULTI-MODAL VIDEO TARGET ACQUISITION  
AND RE-DIRECTION SYSTEM AND METHOD

APPEAL BRIEF

On Appeal from Group Art Unit 2173

Date: September 15, 2005

By:

Gregory L. Thorne  
Attorney for Applicant  
Registration No. 39,398



TABLE OF CONTENTS

	<u>Page</u>
I. REAL PARTY IN INTEREST.....	4
II. RELATED APPEALS AND INTERFERENCES.....	5
III. STATUS OF CLAIMS.....	6
IV. STATUS OF AMENDMENTS.....	7
V. SUMMARY OF THE CLAIMED SUBJECT MATTER...	8
VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL.....	14
VII. ARGUMENTS.....	15
VIII. CONCLUSION.....	26
IX. APPENDIX: THE CLAIMS ON APPEAL.....	27



TABLE OF CASES

<b>In re Oetiker</b> , 977 F.2d 1443, 1445, 24 USPQ2d 1442, 1444 (Fed. Cir. 1992).....	15
<b>In re Piasecki</b> , 745 F.2d 1468, 1472, 223 USPQ 785, 788 (Fed. Cir. 1984).....	15
<b>In re Fine</b> , 837 F.2d 1071, 1074, 5 USPQ2d 1596, 1598 (Fed. Cir. 1988).....	15
<b>In re Lee</b> , 277 F.3d 1338, 1343, 61 USPQ 1430, 1433 (Fed. Cir. 2002), citing <b>McGinley v. Franklin Sports, Inc.</b> , 262 F.3d 1339, 1351-52, 60 USPQ2d 1001, 1008 (Fed. Cir. 2001).....	15
<b>In re Dembiczak</b> , 175 F.3d 994, 999, 50 USPQ2d 1614, 1617.....	15
<b>McElmurry v. Ark. Power &amp; Light Co.</b> , 995 F.2d 1576, 1578, 27 USPQ2d 1129, 1131 (Fed. Cir. 1993) .....	16

I. REAL PARTY IN INTEREST

The real party in interest is the assignee of the present application, Philips Electronics North America Corp., and not the party named in the above caption.

**II. RELATED APPEALS AND INTERFERENCES**

With regard to identifying by number and filing date all other appeals or interferences known to Appellant which will directly effect or be directly affected by or have a bearing on the Board's decision in this appeal, Appellant is not aware of any such appeals or interferences.

**III. STATUS OF CLAIMS**

Claims 1, 2, 4, and 6-15 are presented for examination. All of these claims are pending. Claim 12 is indicated as allowed. Claim 11 is indicated as allowable if amended to be in independent form including all the limitations of the base claim and any intervening claims.

Claims 1, 2, 4, 6-10 and 13-15 stand finally rejected, and form the subject matter of the present appeal.

IV. STATUS OF AMENDMENTS

The Amendment after the Final Office Action containing Claims 1, 2, 4, and 6-15 filed June 3, 2005 has been entered.

**V. SUMMARY of the CLAIMED SUBJECT MATTER**

The present system relates generally to the field of video-camera systems, such as a video conferencing systems and methods, and more particularly to video camera targeting systems and methods (for example as shown in FIG. 1) that locate and acquire targets (e.g., target object 5) using one or more input sensed trigger events, such as user voice input from sound sensor 8 and video input sensing a user's gesture from cameras, such as cameras 1, 44 (e.g., see patent application, page 14, lines 18-20). Another triggering event may include a triggering event from push button 15 (e.g., see patent application, FIG. 1B and the accompanying text on page 17, lines 16-20). An example of a sequence of the current system using user voice input as a triggering event is shown in FIG. 2. An example of a flow chart for the current system using video input sensing a user's gesture as a triggering event is shown in FIG. 3. The above operation corresponds to "sensing a triggering event generated by a human operator" in terms of Claim 1.

The current system also identifies machine sensible characteristics of potential targets (in terms of Claim 1, "receiving additional external information that

characterizes at least one machine-sensible feature of a target"). Examples of machine sensible characteristics include object features that are determined using video input, such as from object camera 2 (e.g., see patent application, FIG. 1, and the accompanying text on page 15, lines 3-5) and object features that are determined by infrared sensor 6 (e.g., see patent application, page 15, lines 5-8). An exemplary flow chart of the present system identifying and storing potential target's machine sensible characteristics is shown in FIG. 7, during acts E-1 and E-2 (e.g., see patent application, page 22, lines 2-13). As described, one camera can both receive an input from the user, e.g., pointing gesture, and acquire a target image (e.g., see patent application, page 19, lines 12-20). Examples of some potential machine sensible characteristics include visual features of targets, such as colors, patterns (e.g., see patent application, page 21, lines 5-7) and identification of targets in commonly recognized terms, e.g., a book, for use as a machine discernable characteristic (e.g., see patent application, page 22, lines 20-21). In addition, unrecognized targets can be compared to recognized targets to identifying similarities

of characteristics and thereby, discern machine sensible characteristics (e.g., see patent application, page 22, line 25 through page 23, line 5), etc.

The present system uses the triggering event together with the machine sensible characteristics to help aim a camera 2 in the direction of the target 5 (in terms of Claim 1, "aiming a camera in response to said sensing and said receiving step", e.g., see patent application, page 6, lines 4-15). The use of these multiple inputs (both a triggering event from the user and machine discernable characteristics of potential targets) has a clear advantage over prior systems in that the present system eliminates false targets while enabling the user to act in a natural manner with little or no training (e.g., see patent application, page 5, line 25 through page 6, line 3).

**A. Example of a System That Falls Within the Scope of  
Claims 1 and 13**

A simple exemplary embodiment that falls within the scope of Claim 1 is instructive of some of its advantages. In the exemplary embodiment, a person points in the direction of a barrel and says the word "barrel". (See, the patent application, page 6, lines 28-29, page 7, lines 5-8 and lines 24-26, and Fig. 1A target object 5.) A first camera of the system recognizes the pointing as a triggering event using the gesture as one input in identifying the target barrel (patent application, page 6, lines 23-27 and page 15, lines 21-24.). In addition, the spoken word "barrel" may be captured by the system (e.g., see, patent application, page 7, lines 5-8). The spoken word "barrel" is a generic name that describes an object and serves as a characterization that may correspond to at least one machine sensible feature of the target barrel.

It is important to note that as recited in Claim 1 the "additional external information" received "characterizes at least one machine-sensible feature of a target ...". Thus, in the above example, processing of the speech in itself (for example, via a speech processor) does not

necessarily mean that the speech "characterizes at least one machine-sensible feature of a target ..." In the example, the speech input received qualifies as "additional external information that characterizes at least one machine-sensible feature of a target" because "barrel" characterizes a target feature that may be captured by a camera and detected by image processing.

Thus, Claim 1 includes a number of advantageous features. For example, receiving additional external information regarding a target substantially simultaneously with the sensing of a triggering event that includes sensing a gesture provides that the additional external information is more reliably identified and more rapidly processed. Also, the additional external information received "characterizes at least one machine sensible feature of a target ..." Such received information may be utilized to great advantage in locating a target. For example, the above-example demonstrates that the external information received may be correlated with the machine sensible feature of the target, resulting in more flexibility and accuracy in locating the target.

**B. Example of a System That Falls Within the Scope of**

**Claim 14**

As indicated in the example above, the gesture included in the triggering event may also be used as one input in identifying the target. Thus, another aspect of the invention includes inputting spatial information to indicate a position of a target where the spatial information includes sensing a gesture indicating a direction of the target. Use of this information together with spoken or other input information about the target provides more accuracy. An exemplary embodiment that includes this aspect is described in the patent application at page 19, lines 1-5 and lines 21-24, referring to FIG. 2, and pointing trajectory 367. Independent Claim 14 includes recitation relating to this aspect.

**VI. GROUND OF REJECTION TO BE REVIEWED ON APPEAL**

The issues in the present matter are:

**A. whether Claims 1, 2, 4, 6, 7, and 13 are patentable over "The IntelliMedia WorkBench A Generic Environment For Multimodal Systems", by Brondsted et al. ("Brondsted") in view of "Combining Audio and Video in Perceptive Spaces", by Wren et al. ("Wren"); or whether Claims 8-10 are patentable over Brondsted in view of Wren, in further view of "Toward Natural Gesture/Speech HCI: A Case Study Of Weather Narration" by Poddar et al. ("Poddar"); or**

**B. whether Claims 14-15 are patentable over Brondsted in view of Wren.**

## VII. ARGUMENTS

### A. Standard for Prima Facie Case of Obviousness

In rejecting claims under 35 U.S.C. § 103, the Examiner bears the initial burden of establishing a **prima facie** case of obviousness. In **re Oetiker**, 977 F.2d 1443, 1445, 24 USPQ2d 1442, 1444 (Fed. Cir. 1992). See also In **re Piasecki**, 745 F.2d 1468, 1472, 223 USPQ 785, 788 (Fed. Cir. 1984). The Examiner can satisfy this burden by showing that some objective teaching in the prior art of knowledge generally available to one of ordinary skill in the art suggests the claimed subject matter. In **re Fine**, 837 F.2d 1071, 1074, 5 USPQ2d 1596, 1598 (Fed. Cir. 1988).

When determining obviousness, "[t]he factual inquiry whether to combine references must be thorough and searching." In **re Lee**, 277 F.3d 1338, 1343, 61 USPQ 1430, 1433 (Fed. Cir. 2002), citing **McGinley v. Franklin Sports, Inc.**, 262 F.3d 1339, 1351-52, 60 USPQ2d 1001, 1008 (Fed. Cir. 2001). "It must be based on objective evidence of record." *Id.* "Broad conclusory statements regarding the teaching of multiple references, standing alone, are not 'evidence.'" In **re Dembiczak**, 175 F.3d 994, 999, 50 USPQ2d 1614, 1617. "Mere denials and conclusory statements,

however, are not sufficient to establish a genuine issue of material fact." *Dembiczak*, 175 F.3d at 1000, 50 USPQ2d at 1617, citing *McElmurry v. Ark. Power & Light Co.*, 995 F.2d 1576, 1578, 27 USPQ2d 1129, 1131 (Fed. Cir. 1993).

B. Rejection of Claims 1, 2, 4, 6, 7 and 13 under 35  
USC §103(a) in view of Brondsted and Wren alone, or in  
further view of Poddar

Claims 1, 2, 4, 6, 7 and 13 are rejected under 35 USC §103(a) as obvious over Brondsted in view of and Wren. The Final Office Action states that all of the elements of the claimed invention are disclosed by Brondsted in view of Wren alone, or in further view of Podder. This assertion is respectfully traversed.

Claim 1 recites in pertinent part "receiving additional external information that characterizes at least one machine-sensible feature of a target..." The "additional external information" recited in Claim 1 is "additional" to sensing a triggering event that includes sensing a gesture indicating a direction of the target.

While it is true that in some exemplary instances, the "additional external information" recitation of Claim 1 may include speech input, for example. However, Claim 1 requires that the additional external information, such as speech input, "characterizes at least one machine-sensible feature of a target", among other recitations.

The current Office Action maintains reliance on the spoken input of Brondsted for purportedly showing the Claim 1 recitation of "receiving additional external information that characterizes at least one machine-sensible feature of a target..." (See, the Office Action, page 3, lines 4-9.) However, the speech input "Show me Hanne's office" (for the Campus Information System) and undefined spoken commands (for the Automatic Pool Trainer), do not characterize a machine-sensible feature of a target in Brondsted.

While it is true that Brondsted may recognize speech input (see, Brondsted, FIG. 2), however, to qualify as "additional external information", such speech input must of course fall within the scope of the pertinent recitations of Claim 1.

The Examiner's Response asserts that Brondsted in view of Wren teaches the Claim 1 recitation of "receiving

additional external information that characterizes at least one machine-sensible feature of a target..." (See, the Office Action, ¶7 at page 12.) However, the ensuing points made in paragraph 7 of the Office Action fail to support that assertion.

For example, the Office Action states that "the [Brondsted Campus Information] system allows the user to ask questions about the location of persons and offices, labs, etc, then the system analyzes the question or the spoken word (via one or more modules, Fig. 2) and outputs the intended output ..." (See, the Office Action, ¶7 at page 13, lines 12-14.) However, noting that one or more system modules analyze the spoken input does not support the Office Action's conclusion that "[t]he [Brondsted Campus Information] system therefore receives additional external information that characterizes at least one machine-sensible feature of a target." (See, the Office Action, ¶7 at page 13, lines 16-17.)

As previously noted, Claim 1 recites that the "additional external information" received "characterizes at least one machine-sensible feature of a target..." The Examiner's assertion that Brondsted processes certain

speech input by itself is not a showing that the speech input "characterizes at least one machine-sensible feature of a target" as recited in Claim 1.

Notoriously absent from the Office Action's response is any assertion, for example, that the spoken words "Hanne" and/or "office" in some way characterize a feature of a target on the campus map of Fig. 1 that may be detected by a machine sensor. It is respectfully noted that Applicants have stressed these points during the prosecution. (See, e.g., Reply submitted on December 3, 2004, page 4, lines 11-14; and prior Appeal Brief submitted on June 7, 2004, in a paragraph spanning pages 4-5.)

The Response to Arguments at page 13, line 18 though page 14, line 10 of the Office Action also attempts to demonstrate that Brondsted teaches the Claim 1 recitation "receiving additional external information that characterizes at least one machine-sensible feature of a target ..." However, this portion of the Response begins by noting that the Brondsted Campus Information system receives, for example, the spoken inquiry "show me Hanne's office" and that "[t]he inquires ... are analyzed and/or compared (via one or more modules, Fig. 2) with the pre-

stored campus information, and the system retrieves and outputs the answer ..." (See, the Office Action, ¶7 at page 13, lines 18-21.) Thus, it appears this is simply a repetition of the point previously made in paragraph 7 of the Office Action that one or more system modules analyze the spoken input, not a showing that the speech input characterizes at least one machine-sensible feature of a target on the campus map.

Focusing further on this portion of the Response to Arguments, it is briefly noted that it also includes a number of additional unsupported or unclear statements.

For example, the Office Action asserts that spoken inquiries such as "show me Hanne's office" are "characteristic or attribute feature of a target". (See, the Office Action, ¶7 at page 13, line 19.) As noted, however, the Office Action fails to show how such spoken input characterizes a machine-sensible feature of a target on the campus map of Brondsted (individual offices are not even visible, or otherwise evident, on the campus map of Fig. 1, nor is such a level of detail evident from the description of Fig. 1 in Section 2.1 of Brondsted. In addition, the Office Action asserts that the way in which

rooms and offices are described in the Domain Model module of Brondsted teaches machine-sensible features of a target (see, the Office Action, ¶7 at page 14, lines 1-5). How this is taught by the Domain Model module is not contained within any portion of Brondsted, and it is also not clear how this assertion is intended to relate to the pertinent recitations of Claim 1.

Wren is only cited in the Office Action for purportedly showing the Claim 1 recitation of "aiming a camera in response to said sensing and said receiving step" (see, the Office Action, ¶4 at pages 3-4). It is noted that the Office Action cites certain speech input of Section 3.3 of Wren for the "receiving step" aspect of the Claim 1 "aiming" recitation. While the Office Action states that the user of Wren "points to a link (target of interest) and says 'there' to load a new URL page" (see, page 3, line 22 through page 4, line 1), no assertion is made that saying the word "there" in this context characterizes a feature of the displayed link and no assertion is made that any such feature would be "machine-sensible". Further, no assertion is made regarding the other spoken words that Wren mentions in the context of the

"City of News" of Section 3.3. Thus, there is also no showing or assertion that Wren provides the Claim 1 recitation of "receiving additional external information that characterizes at least one machine-sensible feature of a target".

Accordingly, the Office Action fails to show that Brondsted in view of Wren discloses or suggests "receiving additional external information that characterizes at least one machine-sensible feature of a target" as required by Claim 1. For at least this reason, Brondsted in view of Wren fails to present a prima facie case of obviousness with respect to Claim 1 at least under MPEP 2143.03. Reconsideration and allowance of independent Claim 1 is respectfully requested. Independent Claim 13 includes recitations that provide analogous distinctions as discussed for Claim 1 and is distinguished from Brondsted in view of Wren for at least analogous reasons. Reconsideration and allowance of Claim 13 is respectfully requested.

Dependent Claims 2, 4, 6, and 7 are also rejected in paragraph 3 of the Office Action as unpatentable over Brondsted in view of Wren. Without conceding the

patentability per se of dependent Claims 2, 4, 6, and 7, they are distinguishable from Brondsted in view of Wren at least by virtue of their dependency on independent Claim 1. Reconsideration and allowance of Claims 2, 4, 6, and 7 is respectfully requested.

Claims 8-10 are rejected in paragraph 4 of the Office Action as unpatentable over Brondsted in view of Wren and further in view of Poddar. Poddar is not cited for curing any of the deficiencies of Brondsted and Wren described above with respect to independent Claim 1. Accordingly, without conceding the patentability per se of dependent Claims 8-10, the Office Action fails to present a prima facie case of obviousness with respect to Claims 8-10 at least by virtue of their dependencies on independent Claim 1. Reconsideration and allowance of Claims 8-10 is respectfully requested.

C. Rejection of Claims 14-15 under 35 USC §103(a) over

Brondsted in view of Wren

Claim 14 is rejected in the Office Action as unpatentable over Brondsted in view of Wren. Claim 14

recites among other things "inputting spatial information to indicate a position of a target", where the spatial information "includes sensing a gesture indicating a direction of said target". Claim 14 also recites "inputting further information about said target" which may, for example, comprise speech input about the target. Claim 14 also recites orienting an instrument with respect to the target to acquire the target in response to both the spatial and further information "to reduce an ambiguity in said position" of the target.

The Office Action points to the last paragraph of page 5 of Wren as purportedly teaching the Claim 14 recitation "to reduce an ambiguity in said position ..." (See, the Office Action, ¶3 at page 8, lines 13-15 and ¶7 at page 15, lines 4-8.) The cited portion of Wren, however, refers to use of visual cues to activate the speech system, as well as use of speech to disambiguate gestures. The Office Action still fails to show at least the Claim 14 recitation of orienting an instrument with respect to the target to acquire the target in response to both spatial and further information "to reduce an ambiguity in said position" of the target..

For at least this reason, the Office Action fails to present a prima facie case of obviousness with respect to Claim 14 at least under MPEP 2143.03. Reconsideration and allowance of independent Claim 14 is respectfully requested.

Dependent Claim 15 is also rejected in paragraph 3 of the Office Action as unpatentable over Brondsted in view of Wren. Without conceding the patentability per se of dependent Claim 15, it is distinguishable from Brondsted in view of Wren at least by virtue of their dependency on independent Claim 14. Reconsideration and allowance of Claim 15 is respectfully requested.

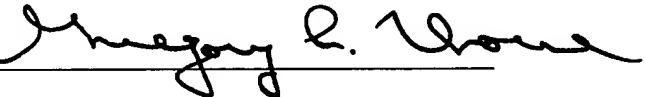
VIII. CONCLUSION

In view of the above analysis, it is respectfully submitted that the referenced teachings, whether taken individually or in combination, fail to anticipate or render obvious the subject matter of any of the present claims. Therefore, reversal of all outstanding grounds of rejection is respectfully solicited.

It is noted that Claim 11 is indicated as allowed and Claim 12 is indicated as allowable.

Early and favorable action is earnestly solicited.

Respectfully submitted,

By 

Gregory L. Thorne, Reg. 39,398  
Attorney for Applicant(s)  
September 15, 2005

THORNE & HALAJIAN, LLP  
Applied Technology Center  
111 West Main Street  
Bay Shore, NY 11706  
Tel: (631) 665-5139  
Fax: (631) 665-5101

**IX. APPENDIX: THE CLAIMS ON APPEAL**

1. A method of locating and displaying an image of a target, the method comprising the steps of:

sensing a triggering event generated by a human operator;

receiving additional external information that characterizes at least one machine-sensible feature of a target, said receiving step occurring substantially simultaneously with said sensing step; and

aiming a camera in response to said sensing and said receiving step, wherein said sensing step includes sensing a gesture indicating a direction of said target.

2. The method of claim 1, wherein said sensing step includes sensing a gesture of a human operator indicating a target.

3. (Canceled)

4. The method of claim 1, wherein said receiving step includes receiving speech from said human operator.

5. (Canceled)

6. The method of claim 4, further including processing said speech for use with at least one machine sensor, said at least one machine sensor and said speech assisting in locating said target.

7. The method of claim 6, wherein said sensing step includes sensing a gesture indicting a direction from said human operator to said target.

8. The method of claim 6, wherein said processing step includes processing said voice information through a look-up table corresponding said speech to search criteria for use with said at least one sensor.

9. The method of claim 8, wherein said look-up table is modifiable.

10. The method of claim 9, wherein said look-up table is modified by receiving information through the on-line global computer network.

11. (Original -- Allowable) The method of claim 9, wherein said look-up table is modified to include an additional voice input and a corresponding search criteria, said added voice input and said corresponding search criteria established by comparing previous association of said added voice input with at least one machine sensible characteristic of at least one correctly identified target associated with said voice input, said machine sensible characteristic being a basis for determining said corresponding search criteria.

12. (Allowed) A method of locating and displaying an image of a target, the method comprising the steps of:

scanning an area within the range of at least one sensor;

identifying potential targets;

storing information concerning machine sensible characteristics and locations of said possible targets;

sensing a triggering event, said triggering event generated by a human operator;

receiving additional external information that characterizes at least one feature of said target, said receiving step occurring substantially simultaneously with said sensing step; and

aiming a camera in response to said sensing, storing and said receiving steps, wherein said sensing step includes sensing a gesture indicating a direction of said target.

13. A method of aiming a camera at a target, comprising the steps of:

inputting an indication of a position of a target;

inputting further information about a machine-sensible characteristic of said target;

aiming a camera at said target in response to said indication and said further information to reduce an error in said aiming, wherein said inputting an indication step includes inputting a gesture indicating a direction of said target.

14. A method of acquiring a target, comprising the steps of:

inputting spatial information to indicate a position of a target;  
inputting further information about said target; and  
orienting an instrument with respect to said target to acquire said target in response to said spatial information and said further information to reduce an ambiguity in said position, wherein said spatial information includes sensing a gesture indicating a direction of said target.

15. A method as in claim 14, wherein said step of orienting includes orienting a camera.